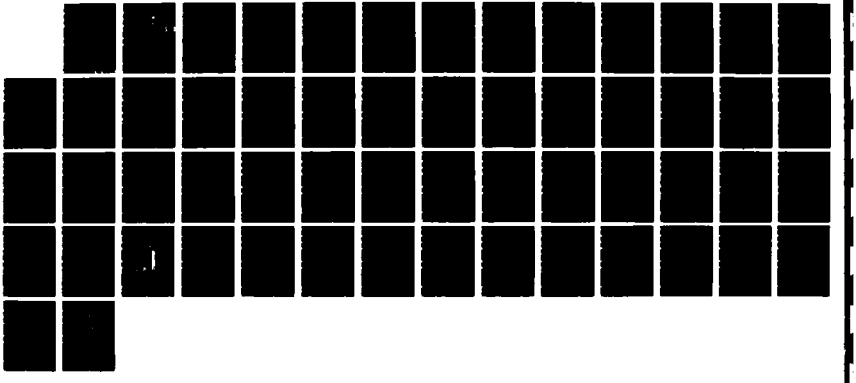
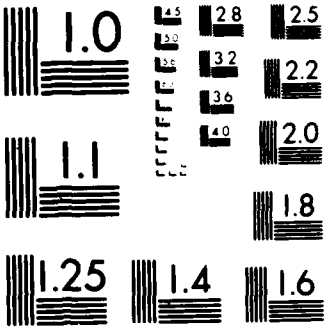
AD-A186 927 A VALIDITY STUDY ON PREDICTORS OF SUCCESS IN RESIDENT 1/1
MASTER'S DEGREE PRO (U) AIR FORCE INST OF TECH
WRIGHT-PATTERSON AFB OH SCHOOL OF SYTT R D WOODWARD
UNCLASSIFIED SEP 87 AFIT/GLM/LSR/87S-86 F/G 5/9 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

A VALIDITY STUDY ON PREDICTORS OF SUCCESS IN RESIDENT MASTER'S DEGREE PROGRAMS AT THE AIR FORCE INSTITUTE OF TECHNOLOGY

THESIS

Rudi D. Woodward
Captain, USAF

AFIT/GLM/LSR/87S-86

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/GLM/LSR/87S-86

A VALIDITY STUDY ON PREDICTORS OF SUCCESS IN RESIDENT MASTER'S DEGREE PROGRAMS AT THE AIR FORCE INSTITUTE OF TECHNOLOGY

THESIS

Rudi D. Woodward
Captain, USAF

AFIT/GLM/LSR/87S-86

Approved for public release; distribution unlimited

The contents of the document are technically accurate, and no sensitive items, detrimental ideas, or deleterious information is contained therein. Furthermore, the views expressed in the document are those of the author and do not necessarily reflect the views of the School of Systems and Logistics, the Air University, the United States Air Force, or the Department of Defense.

AFIT/GLM/LSR/87S-86

A VALIDITY STUDY ON PREDICTORS OF SUCCESS IN RESIDENT MASTER'S DEGREE PROGRAMS AT THE AIR FORCE INSTITUTE OF TECHNOLOGY

THESIS

Presented to the Faculty of the School of Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Masters of Science in Logistics Management

Rudi D. Woodward, B.S.

Captain, USAF

September 1987

Preface

The purpose of this study was to determine the criterion-related validity of predictor variables in measuring graduate grade point averages for resident students at the Air Force Institute of Technology (AFIT), Wright-Patterson AFB, Ohio. Limitations of faculty, facilities, and funds require the Air Force to employ a selective admission policy for its resident master's programs. There is a need for continued research and development of selection models to better the current selection process to help the Air Force better manage its resources.

In performing the research and writing this thesis I have had a great deal of assistance from others. I am very grateful for the constant help and guidance provided by my thesis advisor, Dr. Guy Shane. I also wish to thank my wife Kelly and other family members who stood by me all the way, giving me the support that has made this thesis possible.

Rudi D. Woodward

TABLE OF CONTENTS

Page

LIST OF TABLES

AFIT/GLM/LSR/87S-86

Abstract

This investigation determined the criterion-related validity of 16 predictor variables in measuring graduate grade point averages at the Air Force Institute of Technology (AFIT). Using a sample of 908 Air Force officers who graduated during the period from 1984 to 1986, predictor/criterion relationships were examined and statistical prediction models were developed based on the validity between eligibility criteria and measures of successful models.

The analysis was accomplished by the Stepwise regression method using a .05 level of significance. The results illustrate that 7 of the 16 variables examined were valid predictors of successful performance at AFIT. Prediction models containing these variables were shown to be superior to the present graduate selection process. Prediction models, correlation matrices, and tables of student demographic distributions are presented.

A VALIDITY STUDY ON PREDICTIORS OF SUCCESS IN RESIDENT MASTER'S DEGREE PROGRAMS AT THE AIR FORCE INSTITUTE OF TECHNOLOGY

I. Introduction

The United States Air Force has make a strong commitment to the growth of its people through management and technical education. The Air Force Institute of Technology (AFIT) at Wright Patterson Air Force Base in Ohio is an example of that commitment. At AFIT, both military officers and Department of Defense (DOD) civilian equivalents participate in graduate degree programs leading to master of science and doctoral degrees in management and engineering disciplines. AFIT's in-residence master's degree programs provide Air Force and DOD students with skills necessary for performance at higher echelons in their organizations, thus benefiting the Air Force and the career advancement of the student.

Limitations of faculty, facilities, and funds require the Air Force to employ a selective admission policy for its resident master's programs at AFIT. To maximize its investment, the Air Force only selects officers to attend whose academic and professional job performance indicate a good probability of success in such a demanding environment. To aid in selecting those students likely to succeed, AFIT

has established some eligibility criteria. In general, the minimum mission criteria for the AFIT master's program is a 2.5 undergraduating grade point average (UGPA), on a 4.0 scale, and a standardized test score of at least 1000 on the Graduate Records Examination (GRE), or at least 500 on the Graduate Management Admissions Test (GMAT).

Successful performance in an AFIT resident master's degree program requires completion of all courses with an overall 3.0 average on a scale where A=4.0, B=3.0, C=2.0, D=1.0, and completion of a research thesis on a topic of importance to the DOD. Successful performance in this study will be defined as graduation on time with the required minimum graduate grade point average of 3.0.

Undergraduate grade point averages have been widely used to determine eligibility for graduate and professional schools. But this criterion has become increasingly difficult to interpret due to disparities in grading practices and to the increase of non-traditional degree programs (13:2). These graduate and professional schools are depending more and more on standardized tests such as the GMAT and the GRE to differentiate among student's abilities and chances for success.

These tests allow students from varied backgrounds to be evaluatedon a common ground. Standardized tests can provide much information on the aptitudes and abilities of potential graduate students. It must be assumed, however, that these standardized tests are measuring skills which are

strongly correlated with successful academic performance. These abilities must be reflected by scores that can be ranked, and thereby indicate levels of potential.

These assumptions must be valid for standardized tests to be useful in measuring potential academic performance. Indeed, if the test does not measure skills deemed important and pertinent, then it can be of little use in selecting future students. Criterion-related validity, the correlation between a predictor and a measure of success (criterion), is a measure of the relevance of the test for what it is intended to predict. Graduate admissions departments must have evidence of the standardized test's criterion related validity to insure that the information they receive from test results is relevant to their admissions decisions.

Test users must also be aware of ethical considerations. 'Almost any test can be useful for some functions and in some situations, but even the best test can have damaging consequences if used inappropriately' (1:6). It thus could be argued that a user cannot ethically rely on data from a test until that test's criterion-related validity for a specific purpose has been demonstrated.

There is much published data on validity studies of standardized tests. However, the American Psychological Association contends that 'local collection of evidence on criterion-related validity is frequently more used than published data' (1:18). These and other concerns about the

validity and effectiveness of standardized tests such as the GMAT and the GRE have led many graduate schools to sponsor local validity studies which evaluate standadized tests for their particular purposes. In addition, independent researchers like Furst conclude: 'Each professional school should carry on continued research on the effectiveness of its selection procedures' (11:950).

Practical consideration must be given to the value of data used in the selection process. It is very costly for AFIT to select officers to attend who eventually fail to graduate. In his study in 1983, Van Scotter determined that the average cost associated with sending an officer to AFIT in residence was $82,892.68 for each engineering student and $67,258.66 for each logistic student (24:68). If astudent does not graduate, the figures above can be assumed as total losses to the Air Force, considering an officer could have been selected who would have graduated. It is evident that improvements in the selection process which result in fewer non-graduates could yield significant cost savings.

Eligibility criteria for admission to graduate schools can become outdated. Womer says that 'a test with significant criterion-related validity five or ten years ago may not have the same relationship today' (25:61). Local validity studies can furnish information to aid in the improvement of outdated selection procedures, and more precise prediction models can be developed based on the particular situation.

Problem Statement

Graduate Record Examination (GRE) and Graduate Management Admission Test (GMAT) scores are heavily weighted in the candidate selection process for Air Force Institute of Technology (AFIT) resident master's degree programs. The validity of these test scores and various other predictors of student potential for academic performance used by the registrar's office as selection criteria has not been demonstrated recently.

Until empirical research is accomplished on the criterion-related validity of the present selection process, no basis exists for determining whether AFIT admissions criteria have become invalid or outmoded.

The purpose of this study will be to evaluate the validity of GRE and GMAT test scores and various other indicators as predictors of academic performance in AFIT resident master's degree programs.

Background

Standardized Tests. The use of the GRE and GMAT standardized tests as predictors of performance in graduate programs has been the focus of many studies. Both of these tests have known reliability, and may be used in evaluating academic potential across a wide spectrum of academic disciplines (10:1).

The GRE is an aptitude test designed to predict performance by measuring skills learned over an extended period of

time. The GMAT is a test used primarily by management and business schools. The Educational Testing Service (ETS), which administers the tests, presents data supporting their conclusion that the resulting scores are indeed valid when used to predict graduate performance (10:2).

Tight controls are maintained on the GRE and GMAT to insure standardization in administration, materials, and scoring methods. The ETS uses strict, sound administration procedures to ensure the same specific steps are followed each time the test is given. All versions of the test are exactly the same in appearance, length and format, and each version is reviewed to insure that its content is similar to that of other versions (10:12).

ETS uses scaled and norm scores to report performance on the GRE and the GMAT. 'Scaled score' refers to a basic reference group originally used to establish a scale against which to measure the performance of future examinees. The reference group ETS originally used was a large group of college seniors who took the GRE verbal and quantitative subtests in 1952. The group's mean was made to equal 500, with a set standard deviation of 100. The process is on-going and new reference groups are used to continuously update and validate the tests. ETS statistically manipulates new test score data in order to set the means and standard deviations of new examinees to the same set of parameters. ETS allows for slight errors in measurement, and then states that comparison of test scores between two

or more examinees is a useful and valid measurement i.e., within reliability limits (10:3).

The GRE and the GMAT are divided into various subtests which measure various aptitudes. The GRE consists of verbal, quantitative, and analytical subtests. The verbal and quantitative subtests were first given in 1952 to the original reference group. The analytical subtest was added to the versions of the GRE in 1977, and analytical scores were reported as a separate category in 1978. Each subtest has been carefully designed to measure aptitudes within that category. For example, the analytical subtest measures one's ability to reason, to reach logical and sensible solutions, and to determine the important factors in given situations. The GMAT contains only two subtests, verbal and quantitative.

These standardized tests provide the typical graduate admissions department with easily interpreted quantitative scores. These results will fit easily into a decision criterion formula.

However, subjective measures are much more difficult to interpret. Motivation, drive, professional pride, and various other subjective variables may contribute to one's performance. Subjective evaluations have been less effective than those based on qualitative or statistical methods because of differences in criterion and rater variability (18:178). Travers has shown that the use of test results

has improved the efficiency of many organizations in education as well as other arenas (22:371).

There is a substantial body of published research dealing with the usefulness and effectiveness of the GRE to predict academic success at the graduate level. Research on the GMAT is much more limited, although many of the criticisms and comparisions are similar. The main problem critics have with the GRE lies with low correlations found in some studies examining relationships between the test and the criterion of graduate grade point average (GGPA). Even so, these correlations are most often higher than any other known predictor the researchers have studied. A review of these pertinent studies will be examined in further detail later.

Validity

As defined earlier, validity is the usefulness of a measurement. According to Womer, criterion-related validity, the main method of prediction, is a measure of the strength of the relationship between a test score (such as the GRE) and a future measure of success (such as graduate grade point averages) (25:61). Many schools frequently use criteria such as GGPA and graduation/non-graduation to increase the accuracy with which they select graduate students who are likely to perform successfully (25:61).

The strength of the relationship between these criteria is measured by the Pearson product-moment correlation

(6:65). Positive correlations between predictor and criterion variables allow predictions based on such variables to be more accurate than decisions made at random. Although the ideal situation would be for the predictor variable and the criterion to be perfectly correlated (r=1.00), most validity coefficients are below .60 in actual practice (25:63). Traxler argues that correlations around

for predicting academic performance at the graduate level (23:473).

There are various factors contributing to the low values of validity coefficients. Validity coefficients tend to be low where the range of aptitude levels in a group is narrow. In fact, as ability levels become more similar it becomes harder to differentiate among individuals within the group (8.2). This phenomenon is called restriction in range. As admissions criteria become more stringent, the resulting group of graduate student is much more homogenous than the population as a whole.

The use of other admissions criteria to compensate for low standardized test scores also contributes to lower validity coefficients. If enough students are admitted to graduate school with low test scores because of other valid compensatory factors, then correlations between test scores (GRE or GMAT) and the GGPA will be lower than if admission were based solely on test scores.

Chronbach states that the use of other relevant admission factors in addition to test scores will usually improve the validity coefficients of the prediction or selection model. Factors such as undergraduate grade point average (UGPA) are commonly used as selection criteria for graduate admissions.

Reliability

Tests and other predictor variables must not only be valid, they must be reliable. Cureton says that there can be no meaningful validity without reliability as a prerequisite (9:94). Anastasi states:

> Test reliability indicates the extent to which individual differences in test scores are attributable to 'true' differences in the characteristics under consideration and the extent to which they are attributable to chance errors (2:103).

The Educational Testing Service states that the reliability of both the GRE and the GMAT test is above 90 percent (10:2;12:3). For the purposes of this study, a high degree of reliability in a predictor variable makes it a more credible indicator.

Prediction

There are two types of prediction, clinical and statistical. Thorndike describes clinical prediction as 'a method which assimilates values in a nonlinear manner to permit flexibility, in that any pattern may be obtained and weighted, regardless of its complexity or uniqueness'

(21:201). This method of prediction is highly judgmental and may be based primarily on theory or unique considerations (18:178). Thorndike doubts that judgmental ways of using test scores will be better than the best linear combination of those scores (21:201).

All the literature reviewed in this study involves the other form of prediction - statistical. Historical data from past performance is used to predict future performance using statistical methods (18:178). Using a statistical stepwise regression procedure with samples of data on various predictor variables, it is possible to obtain information about the relative contribution of these variables to the subject criterion of the prediction model. Commonly, all variables entered into the statistical model are arbitrarily assigned weights, even if they have been previously identified as stronger contributors to prediction. Stepwise regression uses a step-by-step process to place predictors into the model in order of their relative contribution. When the model cannot be improved by adding another predictor, the 'best' model has been selected.

Many studies have used statistical prediction to measure the predictive validity of various criteria. Thacker and Williams reviewed 12 studies, 10 of which used GRE scores as the predictor variable and GGPA as the criterion (20:941). They found correlation values which were not statistically significant and could not be used successfully in predictions (20:939). Thacker and Williams

reported that the variability of the GGPA as a criterion variable and the limited range of the sample (sample size was less than 50 in most of these studies) were somewhat responsible for the findings. They also noted that 'the use of other measurement criteria has not consistently yielded improved correlations' (20:939).

Robertson and Nielsen used faculty ratings instead of GGPA as the criterion for success in their study (17:648). They discovered that combining UGPA in college math and science courses with GRE produced a correlation coefficient of .44 at an .05 significance level. This combination of two predictor variables provided a more accurate prediction model than either the GRE or UGPA alone.

Another study examining the predictive value of the GRE was accomplished using GGPA as the criterion of success (7:429). Camp and Clawson obtained a correlation coefficient of .24 at the .01 level of significance for the GRE verbal and GRE quantitative subtest scores combined. They concluded that this correlation was not sufficiently high to be effective in predicting success. However, Brogden would argue that Camp and Clawson as well as other researchers might be hasty in concluding correlations are not high enough. Brogden states that even slight improvements in correlation and prediction can result in more benefits to an organization (6:65).

Cut-off Scores

The use of cut-off scores on tests as a means of differentiating abilities for graduate programs has also been researched. Borg tested the hypothesis that cut-off scores could be used to determine students who were successful or unsuccessful in graduate programs (4:379). He established test cut-off scores for the GRE verbal subtest through statistical means. Successful students were those whose GGPA was greater than or equal to 3.0, and unsuccessful students were those where GGPA fell below 3.0. Using a sample size of 172, Borg found that using the established cut-off score would have eliminated 72 percent of students who were unsuccessful, but would have also denied eligibility to 27 percent, or 21 students who were in fact successful (4:380).

More commonly, several predictor variables are relevant inadmissions decisions for graduate programs. In such cases, cut-off scores may be established for each relevant predictor. One criticism of these multiple cut-off scores is that individuals may be eliminated from consideration if they score below the cut-off on any one test or predictor. Conversely, there is a method that allows for compensation of scores. Multivariate linear models allow for high abilities in one criterion to offset 'ow scores or weaknesses in another. Chronbach contends that multiple cut-off scores should be used only when spcific prerequisites are required and no other abilities can compensate for them (8:437-438).

There is presently no established analytical method for determining and establishing cut-off scores. In addition, Thorndike asserts that the degree of potential success of a student cannot be determined using multiple cut-off scores, and that this method is not useful when the intent is to select the best qualified applicant (21:199). The combined affect of using multiple cut-off scores forms a non-linear selection model, judgmental in nature, and thus not a statistical (acturarial) model. This approach gives the false impression of being a quantitative method.

Other Studies

GRE and GMAT test scores, as discussed earlier, are rarely used alone in determining a student's suitability and chances of success in graduate studies. The relationships between these other variables and the GGPA criterion of success have been investigated. Of these, one of the most common found in a review of the research was UPGA. Livingston and Turner analyzed 189 Educational Testing Service (ETS) validity studies and found that a combination of GRE scores and UGPA scores predicted graduate achievement better than either variable when used alone (15:1).

Breaugh and Mann used discriminant analysis in an attempt to differentiate between graduates and non-graduates of an MBA program using GMAT and UGPA as the primary predictor variables (5:495). Their model improved the

accuracy in predicting graduation from 52 percent (present admissions committee accuracy) to 69 percent (5:496).

Baird completed a study in 1975 which used graduate students' background information to predict relative success in business and law schools (3:942). Using a sample size of over 2000 graduate students, he found that family background and a student's confidence in his abilities were indeed related to success in law and business schools.

Another study investigated the use of a number of predictor variables in predicting success in a graduate psychology program. Using a sample size of 345, Mehrabian reported that the best predictor was the sum of GRE and Miller's Analogy Test (MAT) scores (16:409). More interesting, however, was the fact that the second strongest predictor of graduate success was the use of letters of recommendation (16:410).

VanScotter performed an analysis with various combinations of 13 predictor variables in an attempt to predict successful performance of graduate students at the Air Force Institute of Technology (24:38). His study produced useful predicitive variables, but several years have passed, new graduate programs have been established, and there are yet more possible predictor variables to be evaluated. The validity of current predictor variables and their correlations to academic performance clearly needs to be researched, hence, this study.

Summary

Prediction methods for successful graduate school performance are an important topic for research. Many approaches and many predictor variables have been used in criterion-related validity studies. Researchers identify many promising techniques for prediction, but few follow up studies are attempted. Researchers opt instead to begin anew and not incorporate previous findings or techniques. Thus, a review of the literature leads to an examination of what methods and variables have not worked well in the past in various specific situations, but does not reveal a consensus on what techniques may be useful in a more general application.

Researchers do agree on one point, however, that continuing investigations and empirical research of criterion-related validity are necessary. Reliance on published data to support the use of the the present selection model cannot be justified. There is clearly room for improvement, and the differences in graduate schools and the students they cater to implies the need for local validity research.

The benefits to the Air Force are substantial. If the current selection process can be improved by a prediction model from this research enough to save the cost of even one nongraduate (approximately $75,000), then the effort will have been worthwhile.

RESEARCH HYPOTHESES

1. Standardized test scores such as the Graduate Record Examination and Graduate Management Admissions test are valid predictors of GGPA.

2. Variables such as time since undergraduate degree (TSUD), enlisted years of military service (EYRS), and commissioned years of military service (CYRS), contribute to the prediction accuracy of these selection models.

3. The 'best' prediction model developed in this study could improve the accuracy of AFIT's current selection process.

4. The correlations between GRE tests (predictors) and graduate grade point average (criterion) will vary between the engineering and the logistics master's degree programs.

II. Method

Explanations of Terms and Abbreviations

The variables to be researched in this study, along with their abbreviations, are defined below.

GREV	GRE verbal test score
GREQ	GRE quantitative test score
GREA	GRE analytical test score
GRET	GRE sum of GRE verbal and quantitative scores
GMAV	GMAT verbal subtest score
GMAQ	GMAT quantitative subtest score
GMAT	GMAT composite score
CYRS	commissioned years of service
EYRS	enlisted years of service
UGPA	undergraduate grade point average
GGPA	graduate grade point average
TSUD	time since undergraduate degree
MS	marital status
SEX	gender

Subjects

The subjects (N=908) in this study are past graduates of in-residence AFIT master's degree programs. This study involved relevant personal and biographical data (see above), from students enrolled in the AFIT graduating classes 1984to 1986, inclusive. An indepth survey of the literature on prediction and criterion-related validity of GRE and GMAT standardized tests and other possible predictor variables used to predict academic performance was accomplished. Possible predictor variables for which historical information was available and accessible were identified.

Data Collection

The information on graduates available was in the graduate educational records in the registrar's office at the Air Force Institute of Technology, Wright-Patterson AFB, Ohio. Incomplete or missing data in relevant biographical, predictor, and criterion categories resulted in a reduction in sample size (deemed insignificant because the results were larger than those commonly reported in the literature).

A census was taken of all Air Force officer resident graduate degree records in the registrar's office for classes 1984 to 1986, inclusive. Data on selected variables was manually recorded and later transferred onto computer files for further statistical analyses.

Data Analysis

Stepwise multiple regression was used to calculate predicition models for the data. This technique weights each predictor directly proportional to its correlation with the criterion variable and in inverse proportion to its correlation with all the other predictor variables. The predictor variables with the highest validity and the lowest overlap with the other predictors in each model is assigned the highest weight. Optimum weights are then developed and assigned to each predictor. The resulting multiple correlation coefficient has the highest validity possible for that set of predictor variables (2:180-183).

A comparison of these multiple correlation coefficients produced the 'best' prediction models available based on the data used. The predictors currently being used in the selection process were compared with those of the new predictor models developed in this study to provide empirical data for evaluating present and possible future admissions systems. The results are reported in Chapter 3.

III. Results

Validity of the Predictors

The correlations between GGPA and each of the predictor variables is listed in Table 1. Correlation coefficients were computed for the entire sample of 908. However, because data points were missing from many of the data records, some correlations were based on much smaller sample sizes. A full correlation matrix may be found in Appendix B.

Table 1

Correlations of predictors with GGPA

VARIABLE :	Rank	UGPA	GREV	GREQ
CORRELATION	.010	.266	.225	.292
SAMPLE SIZE	906	906	769	769
SIGNIFICANCE	0.00	0.00	0.00	0.00

VARIABLE :	GREA	GMV	GMQ	GMAT
CORRELATION	.265	.546	.273	.465
SAMPLE SIZE	761	132	132	132
SIGNIFICANCE	0.00	0.00	0.00	0.00

VARIABLE :	EYRS	CYRS	TSUD	GRET
CORRELATION	.047	.019	.001	.306
SAMPLE SIZE	904	906	906	769
SIGNIFICANCE	0.15	0.56	0.98	0.00

Table 1 illustrates that the GRE tests are most highly correlated with GGPA for the AFIT engineering programs.

GMAT tests were most highly correlated with GGPA for AFIT logistics programs. In the recorded sample, there was only one case where both GRE and GMAT scores were reported. This is because engineering candidates are required to take the GRE and systems and logistics students usually take the GMAT. In addition, UGPA also correlated with GGPA at the

(Rank, EYRS, CYRS, and TSUD), none were significantly correlated with GGPA at the .05 level.

Comparing the Correlations

The correlation coefficients shown in Table 1 were derived from a sample containing 21 different master's degree programs. It follows that the results represent a median between the highest and lowest correlations present in any of the individual programs. These correlation coefficients were based on widely varying sample sizes resulting from missing data points in officer educational records. Some of the differences in correlations can be related to instability associated with such variations. The smallest sample size (132) reported in this study was as large or larger than any reported in the researched literature, but was not sufficiently large to permit a breakout of separate graduate programs.

AFIT Admissions Procedures

Shortly after an Air Force officer is commissioned, his educational records are forwarded to AFIT where they will be kept as long as the officer remains on active duty. Evalua-

tors at AFIT review all educational records and forward the names of those officers with above average records to the Air Force Military Personnel Center (MPC).

To be academically eligible an officer must meet certain minimum eligibility criteria. An undergraduate GPA of at least 2.5 in a related field, and either GRE scores of at least 1000 or GMAT scores of 500 or better are usually required. Minimum criteria are specified in Air Force Manual 50-5, Volume I, para 4-15.

As a result of this initial evaluation, officers determined eligible and who have not already formally applied (volunteered) for AFIT admission are 'centrally identified'. Officers who are not identified in such a manner may request an evaluation from AFIT to identify where their academic defiencies exist. Once these deficiences have been corrected by additional course study and acceptable grades, the officer's records will be re-evaluated and updated as eligible at that time.

MPC career managers review the military records of eligible officers forwarded to them. These managers look at officers who have the required job expertise, acceptable performance ratings, and who are eligible for reassignment. Selection folders are prepared on the officers deemed eligible, and sent to the MPC selection board for review. It is doubtful that this part of the selection process is carried out uniformly because each of the career managers has a

different quota to fill and thus operate independently of one another.

The selection board at MPC selects officers to attend AFIT resident master's degree programs. This board consists of senior officers, and unlike a military promotion board, it is very closely associated with the assignment process. Specific guidelines are explained in Air Force Manual 50-5, Volume I, and are adhered to by the selection board.

Validity of the Procedure

VanScotter, who performed intensive research on the validity of AFIT's selection process during the six year period from 1977 to 1982, estimated the validity of the current process at .35, a level of validity which produced a 90.4% on time graduation rate (24:58). Using Taylor-Russell tables (19,576) it can be shown that increasing the validity of a selection model to .65 should increase the on time graduation rate to 99%.

As the selection process is practiced, officers who request evaluation of their eligibility are required to submit GRE or GMAT scores, whereas officers who are 'centrally selected' are commonly evaluated on the basis of UGPA alone. Since the correlation found between UGPA and GGPA in this study was .266, predictions based solely on this one criterion are questionable. This practice establishes a different set of predictors for those who have furnished standardized test scores and those who have not.

The result is a less stringent evaluation process for 'centrally selected' (nonvolunteer) officers than for volunteers, which actually benefits the nonvolunteers.

Best Prediction Models

Stepwise regression was used to develop prediction models using 16 variables identified in Chapter 2. The Stepwise regression process entered and dropped each variable in turn to insure the best combination of predictors was obtained. The process continued until the set of variables with the highest R square was achieved and no other variables met the criteria for entry into the model. The 'best' prediction models were chosen based on a comparison of the R square values of each.

The Stepwise regression program must have at least 2 observations to perform its analysis. There was only one case where an officer's educational record had both GRE and GMAT scores reported completely. Thus, records with GRE scores reported were run separately from those with GMAT scores. This has no significant impact upon the analysis since candidates for engineering programs usually take the GRE and candidates for logistics programs take the GMAT. In the analyzed sample, all records had one or the other of the scores reported.

Table 2

Multiple regression equation
(using cases with GRE)

PREDICTOR	WEIGHT
UGPA	0.14966005
GREQ	0.06080245
GREA	0.03968607

R SQUARE = 0.12626775
SAMPLE SIZE = 759

Table 3

Multiple regression equation
(using cases with GMAT)

PREDICTOR	WEIGHT
UGPA	0.28436360
GMV	1.94127182

R SQUARE = 0.40712556
SAMPLE SIZE = 132

IV. DISCUSSION AND CONCLUSIONS

Review of the Hypotheses

The first hypothesis, that UGPA, GRE scores, and GMAT scores are valid predictors of GGPA can be supported. A review of the correlation matrix shows all these variables are statistically significant at the .05 significance level. The correlations range from .225 for GREV to .545 for GMV. It can be concluded that these variables are valid predictors of GGPA.

The second hypotheses stated that background variables such as time since undergraduate degree, enlisted years of service, and commissioned years of service contribute to the prediction accuracy of selection models. None of these variables had correlations significant at the .05 level. It can be concluded that EYRS, CYRS, and TSUD are not significant predictors of GGPA for AFIT resident master's degree programs.

The third hypothesis, that the 'best' prediction models developed in this study could improve the accuracy of AFIT's current selection process was supported. The use of statistical procedures in analyzing problems of this kind is well supported in the literature. The results were expected to be superior to those derived by the use of judgemental or intuitive means.

The final hypothesis, that correlations between GRE tests and GGPA will vary between the engineering and the logistics master's degree programs, was supported. The differences in correlation coefficients were small, with the highest variation (.16) between GREQ correlations for the two groups. GRET correlations were the most similar between engineering and logistics programs with a difference of only

Conclusions

The selection accuracy of AFIT is better than that of many private institutions. The validity study described in this report has shown ways in which to combine predictor variables to improve that accuracy.

It is not an easy task to select students for graduate school, and no method is 'best' in all situations. Relationships between predictors and the success they predict will vary from one institution to another and will probably change over time. This study has established the validity of two proposed selection models and of seven predictor variables. All of the information on these predictor variables is contained in the educational records of Air Force officers kept in the registrars office at AFIT. These tools are readily accessible and can be used to aid the selection procedure.

The procedures for determining eligibility for AFIT complicate the selection process. Different procedures for selecting volunteers and non-volunteers and the close

association of the MPC selection board with the assignment process hinders the selection of the best possible candidates for graduate school. To work within this environment, certain steps could be taken to make the process more equitable to potential students. Having all officers submit GRE or GMAT scores and eliminating the eligibility for assignment criterion would improve the process.

AFIT has emphasized requirements on submitting test scores, as evidenced by the fact that there were far fewer missing test scores in this study than in a similar study conducted by VanScotter in 1983. This is good for the selection process. Models derived from statistical methods depend on availability of data for significant results. If data are unavailable, then sample sizes are decreased and the model cannot evaluate cases with missing data.

This study is relevant to an important issue in today's Air Force. Decreasing budgets levied by Congress force the Air Force to make the best use of its available resources. The costs involved in selecting students to attend AFIT who will not graduate or in not selecting those who would have is high, and with continued inflation, will continue to increase. Various constraints will no doubt make some of these costs unavoidable. These costs are not always thought of in dollar terms, but they are real, and should be minimized whenever possible.

Appendix A: Correlation Matrices

MATRIX OF CORRELATION COEFFICIENTS

	RANK	UGPA	GREV	GREQ	GREA	GRET	GMV
RANK	1.00	.052	.087	.057	.052	.085	.070
UGPA		1.00	.215	.168	.161	.226	.413
GREV			1.00	.408	.523	.836	
GREQ				1.00	.649	.854	
GREA					1.00	.694	
GRET						1.00	
GMV							1.00

	GMQ	GMAT	EYRS	CYRS	TSUD	GGPA
RANK	.038	.082	.011	.071	.071	.030
UGPA	.324	.404	.051	.049	.057	.266
GREV			.052	.067	.054	.225
GREA			.022	.047	.061	.291
GRET			.043	.067	.068	.306
GMV	.401	.791	.066	.007	.006	.545
GMQ	1.00	.735	.050	.050	.056	.273
GMAT		1.00	.047	.051	.041	.465
EYRS			1.00	.018	.011	.047
CYRS				1.00	.943	.019
TSUD					1.00	.001
GGPA						1.00

CORRELATIONS OF GRE TESTS WITH GGPA
(Engineering Students)

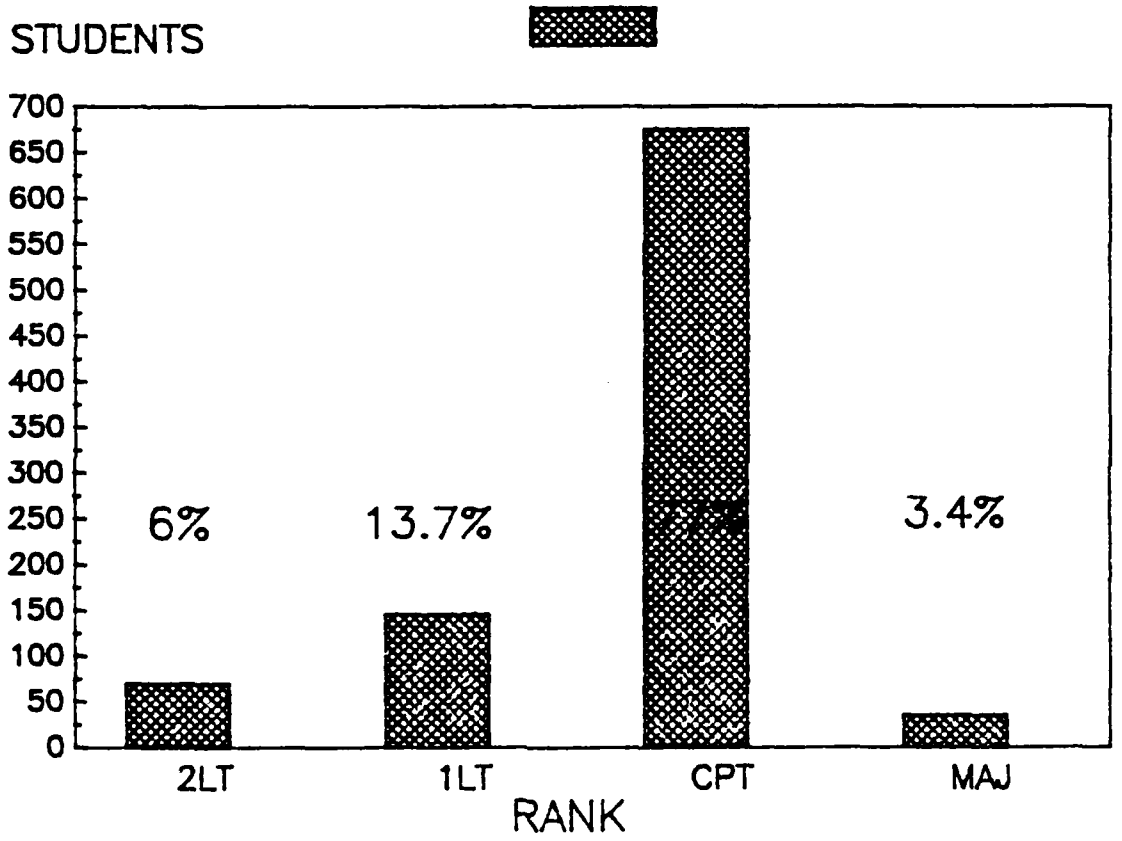
VARIABLE:	GREV	GREQ	GREA	GRET
CORRELATION	.215	.408	.295	.316
SAMPLE SIZE	692	692	681	692
SIGNIFICANCE	0.00	0.00	0.00	0.00

CORRELATIONS OF GRE TESTS WITH GGPA
(Systems and Logistics Students)

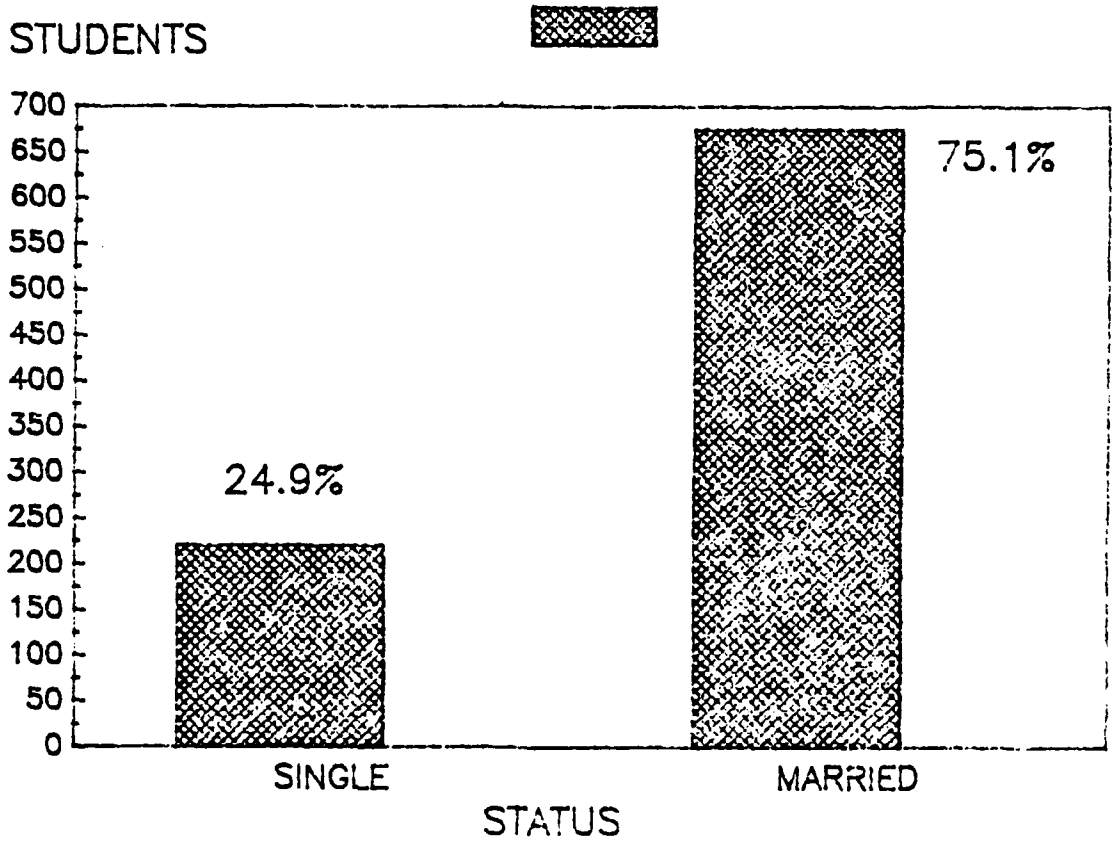
VARIABLE:	GREV	GREQ	GREA	GRET
CORRELATION	.255	.248	.260	.294
SAMPLE SIZE	77	77	69	77
SIGNIFICANCE	0.00	0.00	0.00	0.00

Appendix B: Demographic Information

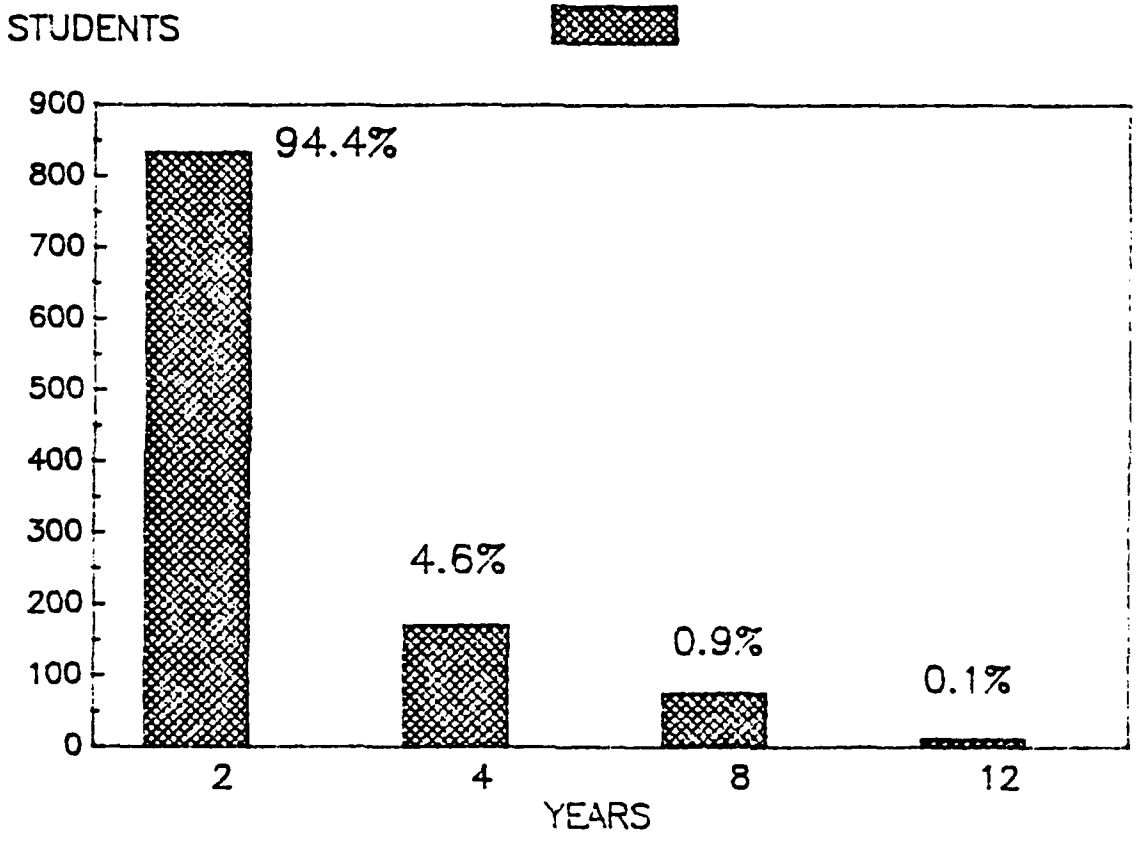
AFIT STUDENT MILITARY RANKS (1984–1986)



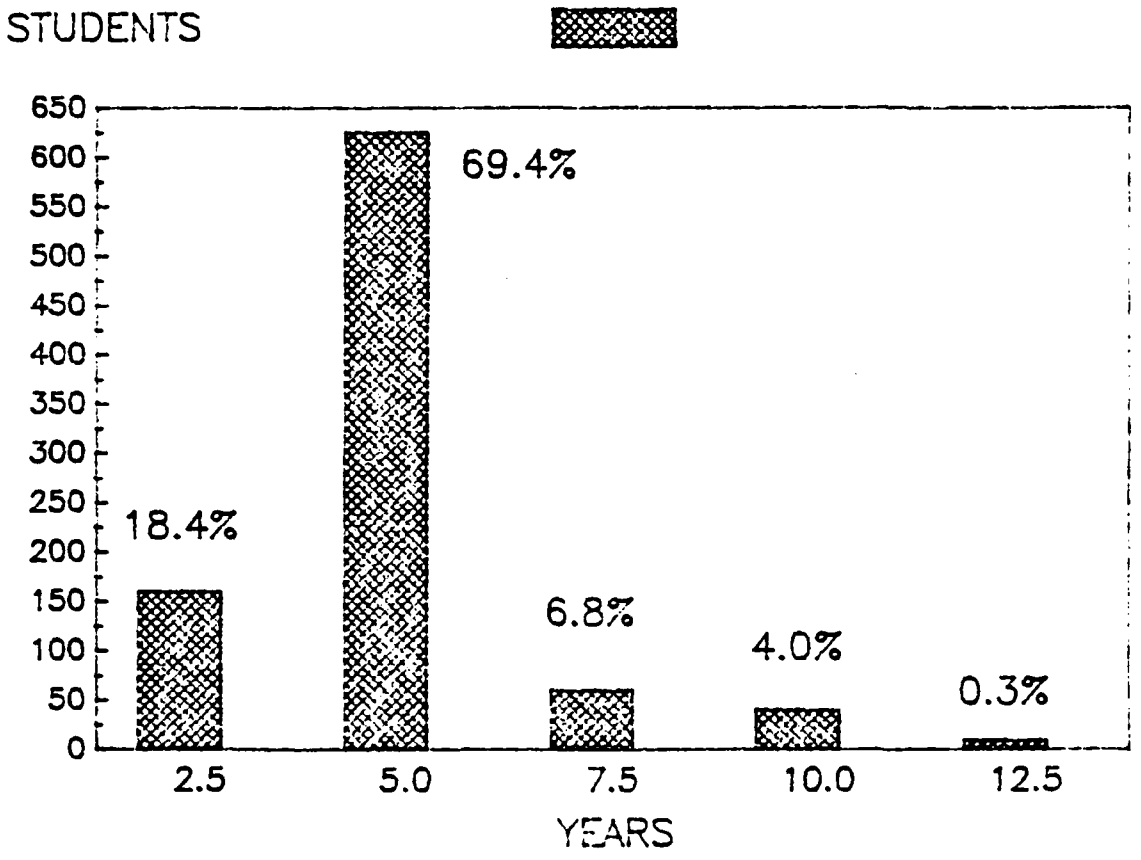
AFIT STUDENT MARITAL STATUS (1984–1986)



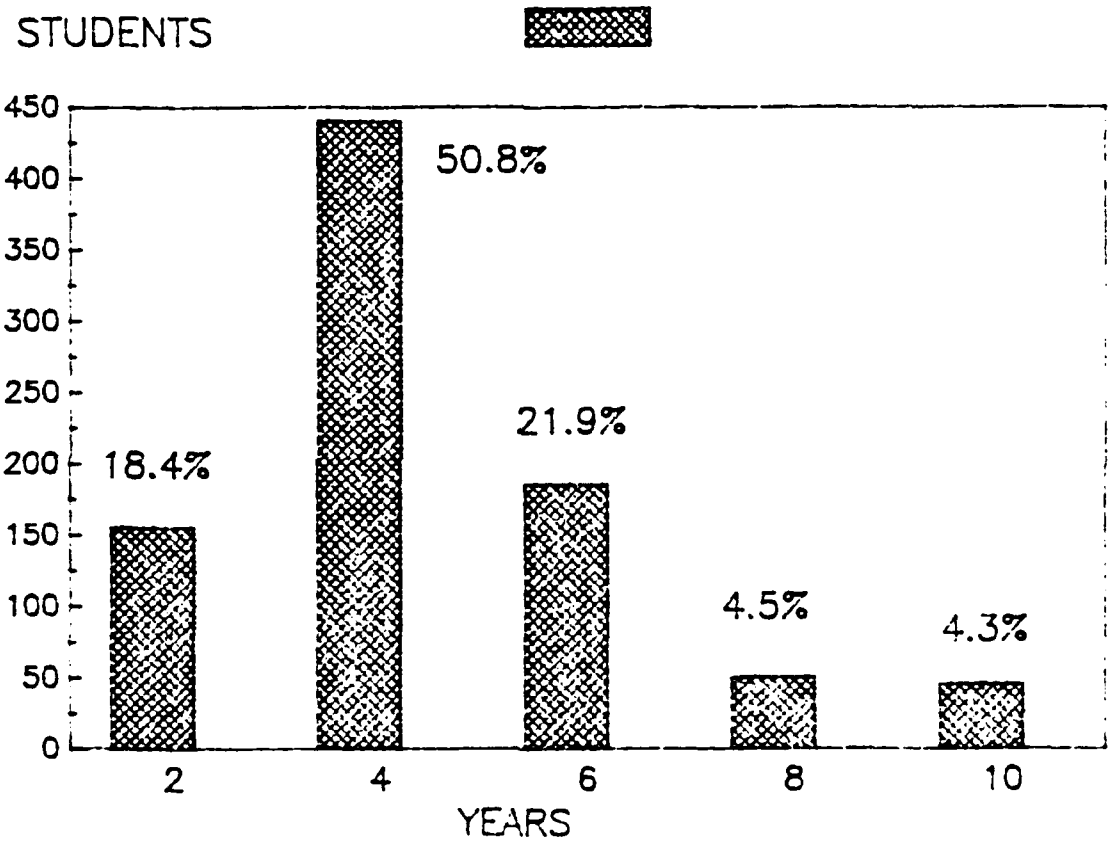
AFIT STUDENT ENLISTED YEARS OF SERVICE (1984–1986)



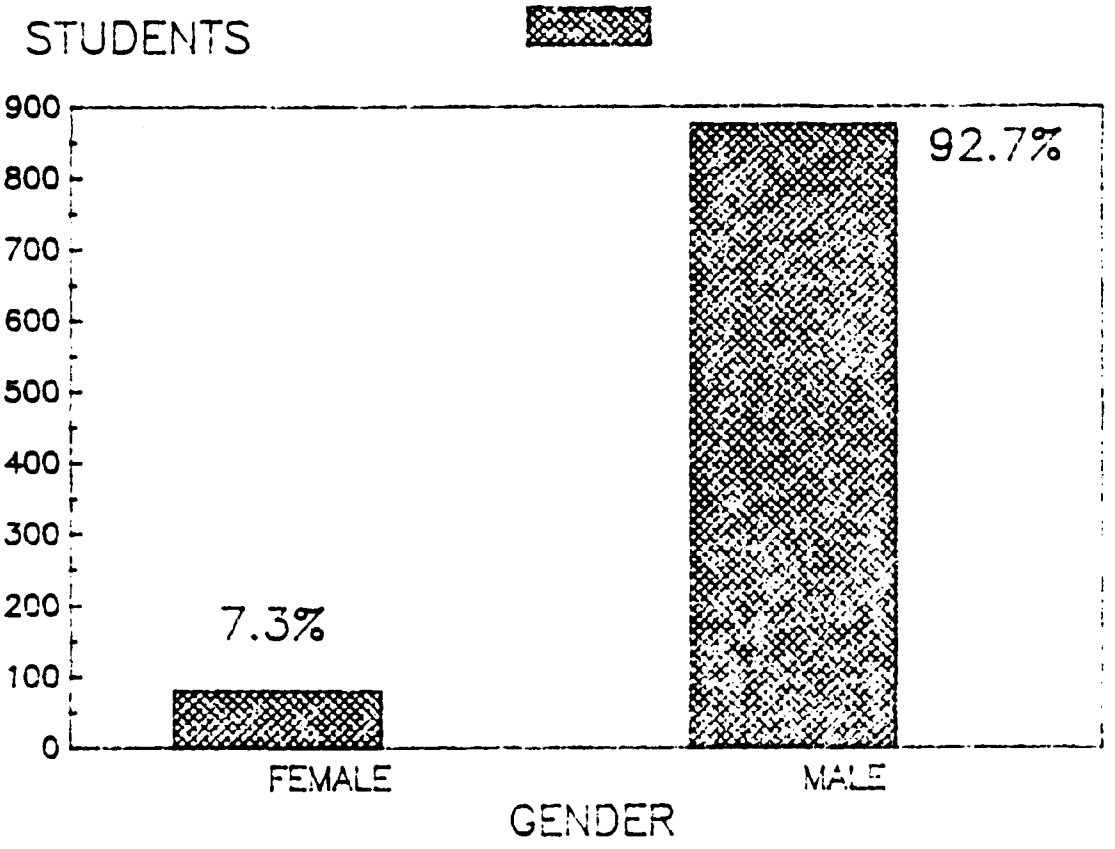
AFIT STUDENT TIME SINCE UNDERGRADUATE DEGREE (1984–1986)



AFIRT STUDENT COMMISSIONED YEARS OF SERVICE (1984–1986)



AFIT STUDENT GENDER
(1984–1986)



Appendix C: Taylor-Russell Tables

PROPORTION OF EMPLOYEES CONSIDERED SATISFACTORY = .70
SELECTION RATIO

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70
.05	.72	.72	.72	.72	.72	.71	.71	.71	.71	.70	.70
.10	.77	.76	.75	.74	.73	.73	.72	.72	.71	.71	.70
.15	.80	.79	.77	.76	.75	.74	.73	.73	.72	.71	.71
.20	.83	.81	.79	.78	.77	.76	.75	.74	.73	.71	.71
.25	.86	.84	.81	.80	.78	.77	.76	.75	.73	.72	.71
.30	.88	.86	.84	.82	.80	.78	.77	.75	.74	.72	.71
.35	.91	.89	.86	.83	.82	.80	.78	.76	.75	.73	.71
.40	.93	.91	.88	.85	.83	.81	.79	.77	.75	.73	.72
.45	.94	.93	.90	.87	.85	.83	.81	.78	.76	.73	.72
.50	.96	.94	.91	.89	.87	.84	.82	.80	.77	.74	.72
.55	.97	.96	.93	.91	.88	.86	.83	.81	.78	.74	.72
.60	.98	.97	.95	.92	.90	.87	.85	.82	.79	.75	.73
.65	.99	.98	.96	.94	.92	.89	.86	.83	.80	.75	.73
.70	1.00	.99	.97	.96	.93	.91	.88	.84	.80	.76	.73
.75	1.00	1.00	.98	.97	.95	.92	.89	.86	.81	.76	.73
.80	1.00	1.00	.99	.98	.97	.94	.91	.87	.82	.77	.73
.85	1.00	1.00	1.00	.99	.98	.96	.93	.89	.84	.77	.74
.90	1.00	1.00	1.00	1.00	.99	.98	.95	.91	.85	.78	.74
.95	1.00	1.00	1.00	1.00	1.00	.99	.98	.94	.86	.78	.74
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.88	.78	.74

PROPORTION OF EMPLOYEES CONSIDERED SATISFACTORY = .30
SELECTION RATIO

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30
.05	.33	.32	.32	.32	.31	.31	.31	.31	.31	.30	.30
.10	.35	.35	.34	.33	.33	.32	.32	.31	.31	.31	.30
.15	.38	.37	.36	.35	.34	.33	.33	.32	.32	.31	.31
.20	.40	.39	.37	.36	.35	.34	.34	.33	.32	.31	.31
.25	.42	.41	.39	.38	.37	.36	.35	.34	.33	.32	.31
.30	.44	.42	.40	.39	.38	.37	.36	.34	.33	.32	.31
.35	.45	.44	.42	.40	.39	.39	.37	.35	.34	.32	.31
.40	.46	.45	.43	.42	.40	.39	.38	.36	.35	.33	.32
.45	.47	.46	.45	.43	.42	.40	.39	.37	.35	.33	.32
.50	.48	.47	.46	.44	.43	.41	.40	.38	.36	.34	.32
.55	.49	.48	.47	.45	.44	.42	.41	.39	.37	.34	.32
.60	.49	.49	.48	.46	.45	.44	.42	.40	.37	.34	.33
.65	1.00	.49	.48	.47	.46	.45	.43	.41	.38	.35	.33
.70	1.00	1.00	.49	.48	.47	.46	.44	.42	.39	.35	.33
.75	1.00	1.00	1.00	.49	.48	.47	.45	.43	.40	.36	.33
.80	1.00	1.00	1.00	1.00	.49	.48	.46	.44	.41	.37	.34
.85	1.00	1.00	1.00	1.00	1.00	.49	.48	.46	.42	.37	.34
.90	1.00	1.00	1.00	1.00	1.00	1.00	.49	.47	.44	.38	.34
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.49	.46	.39	.34
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.39	.34

BIBLIOGRAPHY

1. American Psychological Association, Standards for Educational and Psychological Tests and Manuals, Washington DC : American Psychological Association, 1966.

2. Anastasi, A., Psychological Testing, New York: MacMillan, 1976.

3. Baird, L.L., "Comparitive prediction of first-year graduate and professional school grades in six fields," Educational and Psychological Measurement, 1975, 35, 941-946.

4. Borg, W.R., "GRE aptitude scores as predictors of GPA for graduate students in education," Educational and Psychological Measurement, 1963, 23, (2), 379-382.

5. Breaugh, J.A. & Mann, R.B., "The utility of a discriminant analysis for predicting graduation from a master of business administration program," Educational and Psychological Measurement, 1981, 41, 495-501.

6. Brogden, H.E., "On the interpretation of the correlation coefficient as a measure of predictive efficiency," The Journal of Educational Psychology, 1946, 37, (2), 65-76.

7. Camp, J. & Clawson, T., "The relationship between the graduate records examination aptitude test and grade point average in a master of arts in counseling program," Educational and Psychological Measurement, 1979, 39, 429-431.

8. Chronbach, L.J., Essentials of Psychological Testing, 4th ed, New York : Harper and Row Publishers, 1970.

9. Cureton, E.E., "Validity, reliability, and baloney," Educational and Psychological Measurement, 1950, 10, 94-96.

10. Educational Testing Service, GRE 1981-1982: Guide to the use of the graduate record examinations, Princeton NJ: Educational Testing Service, 1981.

11. Furst, E.J., "Theoretical problems in the selection of students for professional schools," Educational and Psychological Measurement, 1950, (10), 945-952.

12. Hecht, L.W. & Powers, D.E., The predictive validity of preadmission measures in graduate management education: three years of the GMAC validity study service, Princeton NJ: Educational Testing Service, 1982.

13. Knapp, J. & Hamilton, I.B., The effect of non-standard undergraduate assessment and reporting practices on the graduate admissions process, Princeton NJ : Educational Testing Service, 1978.

14. Lin, P. & Humphreys, L.G., "Predictors of academic performance in graduate and professional schools," Applied Psychological Measurement, 1982, 1, (12), 249-257.

15. Livingston, S.A. & Turner, N.J., Effectiveness of the graduate record examinations for predicting first year grades: 1980-1981 summary report of the graduate record examinations validity study service, Princeton NJ: Educational Testing Service, 1982.

16. Mehrabian, A., "Undergraduate ability factors in relationship to graduate performance," Educational and Psychological Measurement, 1969, 29, 409-419.

17. Robertson, M. & Nielsen, W.," The graduate record examination and selection of graduate students," American Psychologist, 1961, (10), 648-650.

18. Sawyer, J.T., "Measurement and prediction, clinical and statistical," Psychological Bulletin, 1966, 66, (3), 178-200.

19. Taylor, H.C. & Russell, J.T., "The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables," Journal of Applied Psychology, 1939, 23, 565-578.

20. Thacker, A.J. & Williams, R.E., "The relationship of the graduate record examination to grade point average and success in graduate school," Educational and Psychological Measurement, 1974, 34, 939-944.

21. Thorndike, R.L., Personnel Selection: Test Measurement and Techniques, New York: John Wiley and Sons, Inc., 1949.

22. Travers, R.M. & Wallace, W.L., "The assessment of the academic aptitude of the graduate student," Educational and Psychological Measurement, 1950, 10, 371-379.

23. Traxler, A.E., "Tests for graduate students," Journal of Higher Education, 1952, 23, 473-482.

24. VanScotter, J.R., Predictors of success in Air Force Institute of Technology resident master's degree programs: a validity study, MS thesis, AFIT/LSSR-83. School of Systems and Logistics, Air Force Institute of Technology (AU),Wright-Patterson AFB OH, September 1983 (AD-A1111111).

25. Womer, F.B., Basic Concepts in Testing, New York: Houghton Mifflin Company, 1968.

Vita

Captain Rudi D. Woodward was born on 29 June 1961 in Oxford, North Carolina. He graduated from high school in 1979 and attended the University of North Carolina on a 4-year Air Force ROTC scholarship. He graduated from there in 1983 with a Bachelor of Science degree in Business Administration and a commission in the USAF. He was assigned to the 3246th Test Wing at the Armanent Division, Eglin AFB, Florida as a munitions accountable supply officer. He served at this station until entering the School of Systems and Logistics, Air Force Institute of Technology, in May 1986.

Permanent address: Route 4, Box 1023
Oxford NC 27565

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A186927

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AFIT/GLM/LSM/87S-86	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION School of Systems and Logistics 6b. OFFICE SYMBOL (*If applicable*) AFIT/LSM	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (*City, State, and ZIP Code*) Air Force Institute of Technology Wright-Patterson AFB OH 45433-6583	7b. ADDRESS (*City, State, and ZIP Code*)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (*If applicable*)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (*City, State, and ZIP Code*)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.
11. TITLE (*Include Security Classification*) See Box 19	
12. PERSONAL AUTHOR(S) Rudi D. Woodward, B.S., Capt, USAF	
13a. TYPE OF REPORT MS Thesis 13b. TIME COVERED FROM _____ TO _____ 14. DATE OF REPORT (*Year, Month, Day*) 1987 September 15. PAGE COUNT 54	
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	09	

18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
Criterion-related validity,
Predictor/criterion relationships

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

Title: A VALIDITY STUDY ON PREDICTORS OF SUCCESS IN RESIDENT MASTER'S DEGREE PROGRAMS AT THE AIR FORCE INSTITUTE OF TECHNOLOGY

Thesis Chairman: Guy S. Shane, PhD
Associate Professor of Organizational Behavior and Management
Dept of Comm and Orgl Sci
School of Systems and Logistics

Approved for public release: IAW AFR 190-1
24 Sept 87
Dean for Research and Professional Development
Air Force Institute of Technology (AU)
Wright-Patterson AFB OH 45433

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Guy S. Shane, PhD	22b. TELEPHONE (*Include Area Code*) (513) 255-2820 22c. OFFICE SYMBOL AFIT/LSR

DD Form 1473, JUN 86 *Previous editions are obsolete*

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Block 19. Abstract

This investigation determined the criterion-related validity of 16 predictor variables in measuring graduate grade point averages at the Air Force Institute of Technology (AFIT). Using a sample of 908 Air Force officers who graduated during the period from 1984 to 1986, predictor/criterion relationships were examined and statistical prediction models were developed based on the validity between eligibility criteria and measures of successful models.

The analysis was accomplished by the Stepwise regression method using a .05 level of significance. The results illustrate that 7 of the 16 variables examined were valid predictors of successful performance at AFIT. Prediction models containing these variables were shown to be superior to the present graduate selection process. Prediction models, correlation matrices, and tables of student demographic distributions are presented.

END

FEB.

1988

DTIC